Expression of human smooth muscle calponin in transgenic mice revealed with a bacterial artificial chromosome

JOSEPH M. MIANO,[1] CHAD M. KITCHEN,[1] JIYUAN CHEN,[1]
KATHLEEN M. MALTBY,[1] LOUISE A. KELLY,[2] HARTMUT WEILER,[3]
RALF KRAHE,[4] LINDA K. ASHWORTH,[5] AND EMILIO GARCIA[5]
[1]Center for Cardiovascular Research, University of Rochester Medical Center,
Rochester, New York 14642; [2]Department of Physiology, Medical College of Wisconsin,
Milwaukee 53226; [3]Blood Center of Southeastern Wisconsin, Milwaukee, Wisconsin 53201-2178;
[4]Division of Human Cancer Genetics, Ohio State University, Columbus, Ohio 43210, and [5]Human
Genome Center, Lawrence Livermore National Laboratory, Livermore, California 94550

Miano, Joseph M., Chad M. Kitchen, Jiyuan Chen, Kathleen M. Maltby, Louise A. Kelly, Hartmut Weiler, Ralf Krahe, Linda K. Ashworth, and Emilio Garcia. Expression of human smooth muscle calponin in transgenic mice revealed with a bacterial artificial chromosome. Am J Physiol Heart Circ Physiol 282: H1793–H1803, 2002. First published January 3, 2002; 10.1152/ajpheart.00875.2001.— Defining regulatory elements governing cell-restricted gene expression can be difficult because cis-elements may reside tens of kilobases away from start site(s) of transcription. Artificial chromosomes, which harbor hundreds of kilobases of genomic DNA, preserve a large sequence landscape containing most, if not all, regulatory elements controlling the expression of a particular gene. Here, we report on the use of a bacterial artificial chromosome (BAC) to begin understanding the in vivo regulation of smooth muscle calponin (SM-Calp). Long and accurate polymerase chain reaction, sequencing, and in silico analyses facilitated the complete sequence annotation of a BAC harboring human SM-Calp (hSM-Calp). RNase protection, in situ hybridization, Western blotting, and immunohistochemistry assays showed the BAC clone faithfully expressed hSM-Calp in both cultured cells and transgenic mice. Moreover, expression of hSM-Calp mirrored that of endogenous mouse SM-Calp suggesting that all cis-regulatory elements governing hSM-Calp expression in vivo were contained within the BAC. These BAC mice represent a new model system in which to systematically assess regulatory elements governing SM-Calp transcription in vivo.

promoter; development; genome

CELLULAR DIFFERENTIATION INVOLVES a carefully orchestrated program of gene transcription resulting in a cell type-specific transcriptome. The completion of the sequencing phase of the Human Genome Project has provided important raw data for the mining of regulatory elements governing the expression profile of a gene (54). This is of particular importance for those genes whose expression is altered in disease states.

Vascular occlusive diseases, for example, are often characterized by the altered expression of a battery of smooth muscle cell (SMC)-restricted genes (40). Understanding how the normal SMC differentiation program is governed represents an important first step toward elucidating the pathways and factors leading to attenuated gene expression in SMC-associated diseases. Accordingly, several groups have characterized the expression and regulation of SMC-restricted genes in a variety of transgenic model systems (8, 11, 14, 15, 21–23, 25, 26, 28, 34, 44, 46, 55). Whereas the majority of these experiments have yielded expected promoter activities based on the endogenous gene's pattern of expression, some of the findings have been unexpected. For example, the endogenous SM22 gene is expressed in all three muscle lineages during development and then becomes restricted to adult vascular and visceral SMC in postnatal life (20). Transgenic mouse studies (15, 21, 34), however, have consistently demonstrated the virtual exclusive activity of the SM22 promoter in arterial SMC of adult animals. In another example, a genomic fragment encompassing the smooth muscle myosin heavy chain promoter and part of the first intron failed to direct expression of a reporter gene in vascular SMC of the lung, head, and neck (26). Together, these reports support the concept of multiple, nonoverlapping regulatory elements in and around gene loci that are necessary for the complete expression profile of a gene.

An important outgrowth of the Human Genome Project that has assisted investigators in defining the boundaries of genomic DNA necessary for the complete expression of a particular gene has been the development of artificial chromosomes (4, 50). Several studies have already exploited these large-capacity cloning vectors in the generation of transgenic animals to characterize DNA sequences necessary for a gene's appropriate pattern of expression. Initial studies using yeast

---

artificial chromosomes showed the appropriate pattern of expression for the human β-globin and human tyrosinase genes (42, 49). Subsequent reports using either yeast artificial chromosomes, bacterial artificial chromosomes (BACs), or P1 phage artificial chromosomes revealed correct spatiotemporal expression of the skeletal muscle regulatory gene myf-5 (57), the existence of an intestinal-specific enhancer located 55-kb upstream of the human apolipoprotein B gene (36, 37), expression of human desmin in all three muscle lineages (45), a repressor of insulin-like growth factor-2 transcription located 40-kb downstream of the gene (1), and juxtaglomerular-restricted expression of the human renin gene (51). Whereas these data demonstrate the power of using artificial chromosomes to capture distal regulatory elements, there have been instances where such an approach has failed to completely recapitulate an endogenous gene's expression profile (19).

Smooth muscle calponin (SM-Calp) is a tightly restricted gene expressed transiently in the embryonic heart, before becoming restricted to all SMC-containing tissues (31, 48). In vitro transient transfection assays have shown that SM-Calp promoter constructs display promiscuous activity when introduced into non-SMC types (16, 31). On the other hand, repeated attempts to recapitulate the endogenous SM-Calp's pattern of expression in transgenic mice using conventional lacZ reporter constructs have shown either no expression whatsoever or ectopic expression due to position effects (unpublished observations, J. M. Miano). To circumvent these difficulties, we have adopted a strategy of using a BAC harboring the entire human SM (hSM)-Calp locus to assess expression of SM-Calp in an in vitro model system of cellular differentiation as well as developing and postnatal transgenic mice. In this report, we describe the initial characterization of this BAC as well as its appropriately directed expression of the hSM-Calp transgene. These studies constitute an important foundation for the discovery of all regulatory elements controlling SM-Calp gene expression in vivo.

MATERIALS AND METHODS

BAC screening and characterization. A 656-bp polymerase chain reaction (PCR) fragment to the hSM-Calp cDNA (forward primer, GCATGGAGCACTGCGACA and reverse primer, CACTGTCACATCCACATAGTATG) was used to screen a ×4 to ×5 high-density arrayed human genomic BAC library according to a commercial protocol (Research Genetics; Huntsville, AL). Two BACs were obtained and subsequently analyzed with PCR primer pairs to the 5′ and 3′ ends of the hSM-Calp transcription unit. One of the BACs (BC-100899) contained the complete hSM-Calp locus and was subsequently used for all studies. BAC DNA was transformed into competent bacteria (DH10Br) and grown in Luria-Bertani Medium containing 12.5 μg/ml chloramphenicol. DNA purification was by alkaline lysis. The size of BC-100899 was estimated by restriction digestion and pulse-field gel electrophoresis (data not shown). The ends of the BAC clone were sequenced and a large-scale *EcoRI* restriction map generated (see clone BC-100899 at http://bbrp.llnl.gov/rmap).
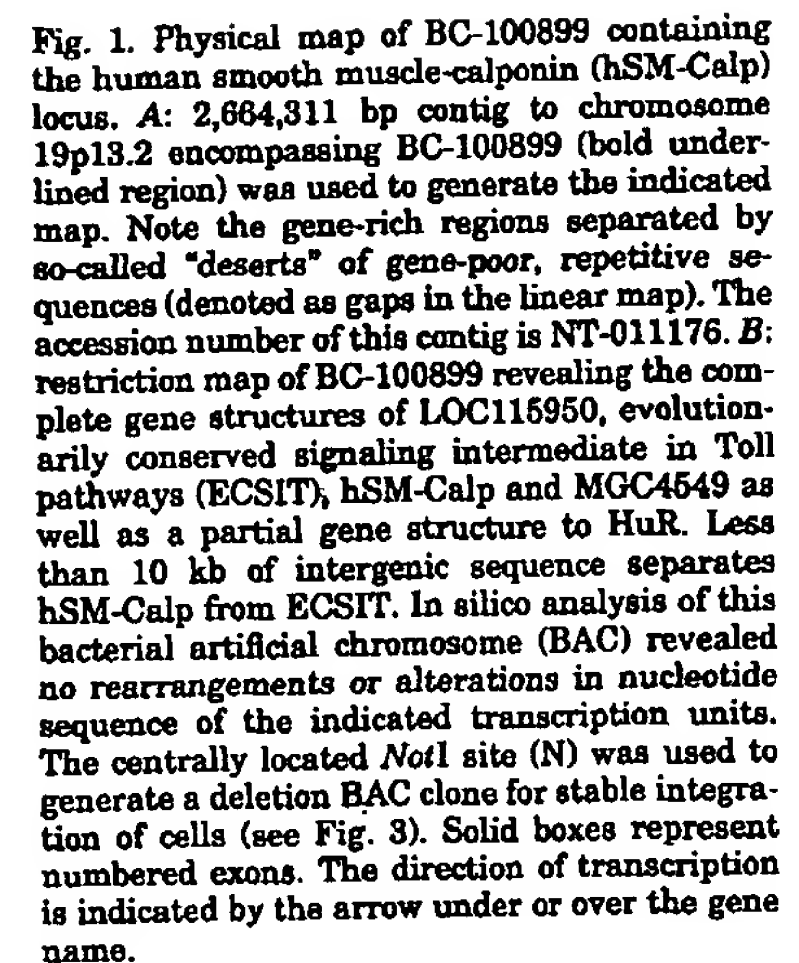
Long and accurate PCR (TaKaRa, Panvera; Madison, WI) was used to ascertain the position of the SM-Calp locus relative to the BAC ends. In addition, we used both long and accurate PCR as well as the NCBI workstation to define the intron-exon boundaries of each transcription unit within BC-100899 (see Fig. 1B).

Generation of stable cell lines. The BC₃H1 and L6 myoblast cell lines (American Type Culture Collection; Manassas, VA) were used to stably integrate BC-100899 and assess its directed expression of the hSM-Calp transgene in vitro. We chose the BC₃H1 cell line as a model system for studying the BAC for two reasons. First, BC₃H1 are the only known cells that reversibly express SM-Calp, making them particularly valuable for studying the potential dynamic expression of a transgene such as SM-Calp (33). Such regulated expression, if conserved in hSM-Calp, would suggest evolutionarily conserved *cis*-elements and *trans*-acting factors whose identification and characterization could be easily carried out in such a cell line. Second, BC₃H1 cells are of mouse origin (31), thus facilitating species homogeneity in the evaluation of the BAC clone in cell culture and the mouse.

Purified BAC DNA (~100 μg) was mixed with 1 μg of a neomycin-containing plasmid (pCI-Neo, Promega; Madison, WI) using lipofectamine (without Plus Reagent). DNA complexes were applied to cells in serum-free Dulbecco's modified Eagle's medium without antibiotics for 12–16 h. Complexes were then aspirated and the cells incubated in serum-containing medium for 24 h. Cells were then trypsinized and replated at varying dilutions (1:20–1:400) in complete medium containing either 600 μg/ml (BC₃H1) or 900 μg/ml (L6) G418. Resistant clones were expanded for PCR genotyping using the primers to hSM-Calp listed above as well as internal control primers to the ZP3 gene (forward, ACGCTCTA-CATCACCTGCCA, reverse, CACTGGGAAGAGACACTCAG). At least three independent clones per cell line were analyzed for expression of the hSM-Calp transgene by RT-PCR and RNase protection assays (see *Generation of transgenic mice*). A deletion construct was similarly tested by cutting the full-length BAC with *Not*I (see Fig. 1B), by religating the fragment containing the hSM-Calp transcription unit into the pBeloBAC vector, and by electroporating DH10Br competent bacteria.

Generation of transgenic mice. Linearized BC-100899 DNA was microinjected into the pronucleus of strain FVB-fertilized oocytes at the Medical College of Wisconsin and Columbia Transgenic Core Facilities. Tail snips from potential founders were obtained and digested at 53°C in lysis buffer (10 mM Tris·HCl, pH 8.0; 25 mM EDTA, pH 8.0; 100 mM NaCl, 1% sodium dodecyl sulfate, and 200 μg/ml proteinase K) for genomic DNA isolation. Initial genotyping was done by PCR using the hSM-Calp and ZP3 primers described above. Southern blotting and DNA hybridizations were also performed to estimate BAC copy number and structural integrity of the BAC clone (data not shown). PCR analysis of the BAC clone invariably revealed loss of one or both ends of the BAC, which had no bearing on the expression of the BAC transgenes. Both transgenic lines reported here have been maintained stably on the FVB background for four generations.

RNase protection assays. Transgene expression analysis required the isolation of total RNA from cells and animal tissues using the acid-phenol guanidinium isothiocyanate procedure (5). As a first estimate of hSM-Calp expression, we performed RT-PCR using human-specific primers flanking two introns of the hSM-Calp cDNA (forward, CGGCAACT-TCATCAACGCCAT and reverse, CACTGTCACATCCACAT-AGTATG). In addition, we analyzed the expression of the

Fig. 1. Physical map of BC-100899 containing the human smooth muscle-calponin (hSM-Calp) locus. *A*: 2,664,311 bp contig to chromosome 19p13.2 encompassing BC-100899 (bold underlined region) was used to generate the indicated map. Note the gene-rich regions separated by so-called "deserts" of gene-poor, repetitive sequences (denoted as gaps in the linear map). The accession number of this contig is NT-011176. *B*: restriction map of BC-100899 revealing the complete gene structures of LOC115950, evolutionarily conserved signaling intermediate in Toll pathways (ECSIT), hSM-Calp and MGC4549 as well as a partial gene structure to HuR. Less than 10 kb of intergenic sequence separates hSM-Calp from ECSIT. In silico analysis of this bacterial artificial chromosome (BAC) revealed no rearrangements or alterations in nucleotide sequence of the indicated transcription units. The centrally located *Not*I site (N) was used to generate a deletion BAC clone for stable integration of cells (see Fig. 3). Solid boxes represent numbered exons. The direction of transcription is indicated by the arrow under or over the gene name.

endogenous mouse SM-Calp gene (forward, ATGTCTTCTG-CACATTTTAAC, reverse, TCAATCCACTCTCTCAGCTCC), human ECSIT (forward, CTGGAGCAGATGGAGAACCAC and reverse, TAGATGGTGACCCTGGCACT), mouse ECSIT (forward, CTGGAGCAGATGGAACGACAC and reverse, CCA-TGCCTGTCGAGTCACTGG), human MGC4549 (forward, GCTCATCCACCTGCAGACATG and reverse, CACGGTA-CAAGAGATGACTCCG), and human LOC115950 (forward, ATGGCAGCAGGCAGCGGTG and reverse, GCACCGTTGT-TGCCAGCTC). As a control, we also examined the expression of glyceraldehyde phosphate dehydrogenase (forward, GCCAAAAGGGTCATCATCTC and reverse, GGCCATCCA-CAGTGTTGT). Total RNA (~3 μg) was reverse transcribed at 37°C for 1 h with a First Strand cDNA Synthesis Kit (Amersham Pharmacia Biotech), as described in the manufacturer's protocol. After reverse transcription, the cDNA templates were individually amplified with the above strand-specific primers. Products were resolved in a 1% agarose gel and visualized by ethidium bromide staining.

To assess the adult tissue distribution of hSM-Calp more rigorously, we performed RNase protection assays with a 319 NT human-specific riboprobe (32) labeled with [^{32}P]UTP as per the manufacturer's instructions (Ambion; Austin, TX). Samples of tissue RNA from two independent transgenic mouse lines were hybridized with labeled riboprobe, RNase A/T digested, and then resolved in a 6% denaturing polyacrylamide gel. Signal intensities were detected by autoradiography (X-OMAT film, Eastman Kodak; Rochester, NY).

Western blotting. Cell monolayers were washed twice with cold phosphate-buffered saline and then scraped in extraction buffer containing 55 mM Tris·HCl (pH 6.8), 10% glycerol, 2% sodium dodecyl sulfate, 10 mM dithiothreitol, 0.5 mM EDTA, 100 μg/ml phenylmethylsulfonyl fluoride, and 1 μg/ml each of pepstatin A, leupeptin, and aprotinin. Alternatively, snap-frozen tissues from transgenic or nontransgenic mice were homogenized in the same buffer and spun at 10,000 g for 20 min for supernatant isolation. Samples of cell or tissue lysates were sheared through a 23-gauge needle and then boiled for 5 min. Protein concentration was determined by the bicinchoninic acid assay (Pierce). Initially, samples of protein extracts (25–50 μg) were resolved through 10% polyacrylamide gel and stained with Coomassie blue to confirm sample integrity and equal loading. Parallel samples were then electroblotted to nitrocellulose (Bio-Rad) and processed for Western blotting as described (31). The SM-Calp antibody (clone hCP, Sigma) was used at a dilution of 1:2,500. As a control, we incubated blots with an antibody to the SH2-adapter protein, Grb2 (c-7, Santa Cruz), or β-tubulin (clone

5H1, PharMingen). Specific immunoreactive proteins were revealed on X-ray film with enhanced chemiluminescence reagents (Amersham Pharmacia Biotech; Piscataway, NJ).

Immunohistochemistry. To assess spatial expression of the hSM-Calp protein, sections of stomach, uterus, and liver were generated from transgenic and nontransgenic mice for immunohistochemistry using a mouse anti-hSM-Calp monoclonal antibody (DAKO; Carpinteria, CA). Antigen retrieval was achieved by microwave heating tissue sections in a solution of 0.1 M citrate buffer (pH 6.3, Zymed; San Francisco, CA) for 10 min. Sections were preblocked for 30 min in normal horse serum, washed and incubated overnight at 4°C with the primary SM-Calp antibody (1:100 dilution). Avidin-Biotin Complex reagents (Vector Laboratories, Burlingame, CA) were then used to reveal hSM-Calp protein with the use of 3-amino-9-ethylcarbazole as substrate.

In situ hybridization. Sections (5 μm) of transgenic and nontransgenic adult tissues and developmentally staged mouse embryos were used to define the spatial and temporal pattern of hSM-Calp mRNA expression. In preliminary studies, we tested several hSM-Calp riboprobes for any evidence of cross-hybridization to the endogenous mouse SM-Calp transcript. The data shown in this report made use of a 3' untranslated riboprobe (260 bp) that was generated with PCR primers to the hSM-Calp cDNA (forward, GCTA-CAGGGTCCAACATAGAAC, reverse, CACTGTCACATCCA-CATAGTATG). This riboprobe is 77% homologous to the corresponding mouse sequence but yielded some cross-hybridization to the endogenous mouse SM-Calp mRNA in developing embryos. Sense and anti-sense riboprobes were linearized and used in vitro transcription assays to generate ^{33}P-labeled riboprobes (MaxiScript, Ambion). Paraffin-embedded, paraformaldehyde-fixed tissue sections were deparaffinized, rehydrated, and treated with 5 μg/ml proteinase K at 37°C for 6 min. Hybridization with ~3 × 10^7 counts/min of probe per milliliter of hybridization solution was performed overnight at 52°C in a humidified chamber. Slides were then washed to remove nonspecifically bound probe, treated with 20 μg/ml RNase A at 37°C for 30 min, dehydrated, air-dried, and dipped in emulsion (Kodak NTB2). After 1 wk, slides were developed in Kodak D19 developer, fixed, and counterstained with Mayer's hematoxylin. Darkfield images were generated with a Polaroid digital camera (DMC-1) using image-capture software (DMC version 2.0) and processed in Adobe Photoshop.

RESULTS

Mapping of transcription units in BC-100899. The hSM-Calp locus maps to chromosome 19p13.2 (27, 32). Figure 1A shows a schematic of a >2.6 Mb map encompassing BC-100899 on chromosome 19p13.2. Chromosome 19 has the highest density of both genes (≈26 genes per Mb DNA) and repetitive sequence in the human genome (13, 53). Both points are illustrated in this large contig with BC-100899 alone harboring five transcription units (Fig. 1A). A higher resolution map of the 103,424 nt BC-100899 is provided in Fig. 1B. As reported previously (32), hSM-Calp comprises 7 exons spanning 11,473 nt. The hSM-Calp locus in BC-100899 is flanked by 9,194 nt of 3' sequence and 82,879 nt of 5' sequence. A divergently transcribed, EST-supported gene comprising 5 exons over 2,421 nt of DNA (MGC4549, accession number NM-032377) resides at the 3' end of BC-100899 immediately after hSM-Calp (Fig. 1B). Three transcription units were discovered 5'

of hSM-Calp and each of these initiates transcription on the opposite strand of DNA. Human evolutionarily conserved signaling intermediate in Toll pathways (ECSIT) (17) is transcribed only 9,678 nt away from the start sites of hSM-Calp (Fig. 1B). The ECSIT transcription unit comprises 8 exons spanning 23,220 nt. A recently predicted mRNA having a zinc fingerlike domain (LOC115950, accession number XM-057078) and consisting of 9 exons over 22,390 nt practically overlaps with ECSIT. Finally, a portion of the HuR transcription unit (24) encompassing 5 exons spans nearly 25 kb of the remaining 5' sequence of BC-100899 (Fig. 1B).

Regulated expression of hSM-Calp in BC3H1 cells. Recently, we (33) reported downregulated expression of SM-Calp in BC3H1 cells induced to differentiate into a skeletal muscle-like lineage, a finding consistent with the absence of SM-Calp expression in developing somites and mature skeletal muscle. To determine whether hSM-Calp displays a similar regulatory pattern of expression, we generated BC3H1 cell lines transfected stably with the BC-100899 clone. Such cells express hSM-Calp mRNA in a growing state (Fig. 2A). In a manner identical to the endogenous mouse SM-Calp gene (33), hSM-Calp transcripts decline on cell differentiation and the adoption of a skeletal muscle-like fate (Fig. 2A). We have observed a similar pattern of mRNA expression in a total of three inde-
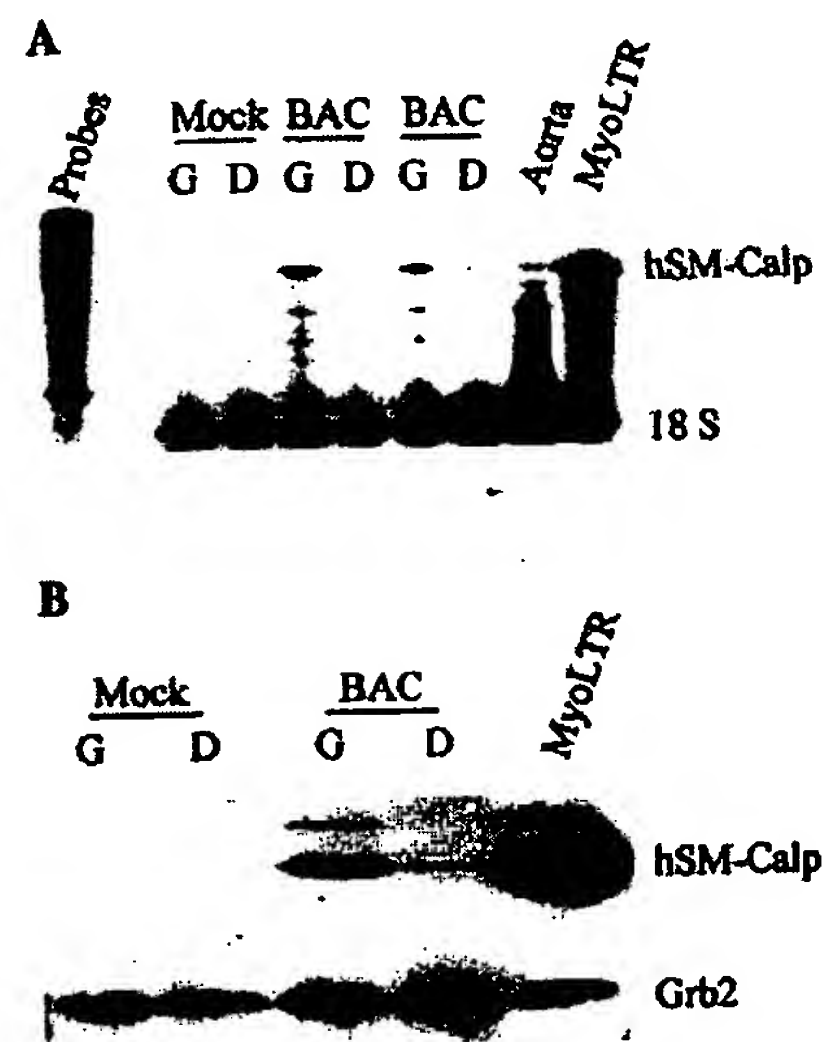

Fig. 2. Expression of hSM-Calp in BAC-transfected BC3H1 cells. Cells were stably transfected with BC-100899 and then analyzed for expression of the hSM-Calp transgene by RNase protection assay (A) or Western blotting (B). Note that both hSM-Calp mRNA and protein were present in growing (G) BC3H1 cells (but not mock transfectants) only to decrease on differentiation (D). Total RNA or protein from human aorta and the human uterine smooth muscle cell (SMC) line (MyoLTR) serve as positive controls for hSM-Calp. Equivalent amounts of total RNA and protein were verified with 18 S and Grb2, respectively.
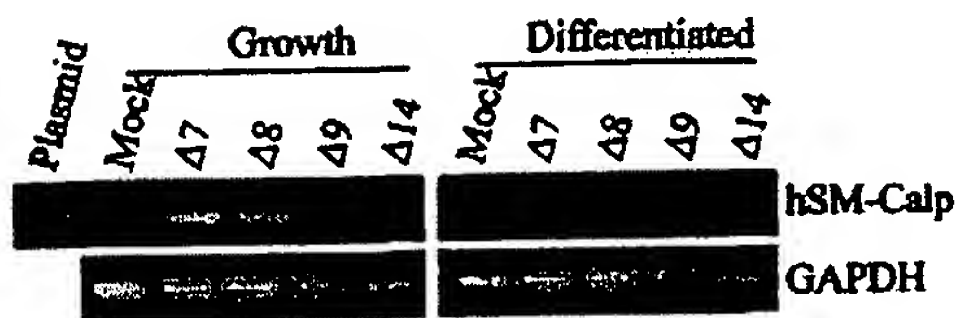
Fig. 3. Expression of deletion mutant BAC in stable BC₃H1 cells. Several independent cell lines (Δ7–9, and 14) carrying the same deletion of BC-100899 (5′ boundary of deletion is at the novel NotI site depicted in Fig. 1B) were selected in G418 medium and their total RNA isolated for RT-PCR analysis. Top: three of the four deletion mutant lines display a prominent signal in growing cells, which decreases when cells are induced to differentiate. In contrast, mock-transfected cells and line Δ14 did not express the hSM-Calp transgene (Southern blot analysis of line Δ14 revealed a different restriction digestion pattern of the BAC clone, suggesting the clone was altered in some manner, data not shown). Bottom: glyceraldehyde-3-phosphate dehydrogenase control product, which is equivalent across all lanes.

pendent BC₃H1 stable cell lines (data not shown). Importantly, stably transfected L6 myoblasts, which do not express endogenous SM-Calp (31), rarely (one line out of >10) exhibited positive expression for the hSM-Calp transgene (data not shown). Western blotting studies demonstrated an immunoreactive band corresponding to hSM-Calp protein that was also downregulated on differentiation of BC₃H1 cells (Fig. 2B). Deletion of the full-length BAC to a NotI site (see Fig. 1B) revealed appropriate expression regulation of hSM-Calp mRNA (Fig. 3), suggesting that the regulatory elements for SM-Calp expression are contained within about 50 kb of genomic sequence. Collectively, these findings demonstrate the transcriptional and translational competence of the hSM-Calp transgene in BC₃H1 cells. Moreover, the hSM-Calp transgene appears to behave in the same manner as the endogenous mouse SM-Calp gene suggesting that the human

locus contains the same BC₃H1 differentiation-responsive regulatory elements as the endogenous mouse SM-Calp locus.

hSM-Calp is restricted to SMC-rich tissues in adult BAC transgenic mice. We generated two independent transgenic mouse lines (each with two copies of BC-100899 as assessed by Southern blotting, data not shown) to evaluate expression of the hSM-Calp transgene in adult tissues. As shown in Fig. 4, mRNA expression of hSM-Calp is essentially identical to that reported previously (Fig. 1A in Ref. 31) with bladder, intestine, stomach, and uterus showing the highest steady-state transcript level. No hybridization signal was observed in nontransgenic bladder and uterus (Fig. 4). Levels of endogenous SM-Calp mRNA did not appear to be affected by the hSM-Calp locus, suggesting that there is no significant promoter competition between the endogenous mouse gene and the hSM-Calp transgene (data not shown). Consistent with a previous report in mice (17), human ECSIT, which is <10 kb away from the hSM-Calp locus, is widely expressed in adult tissues (Fig. 5). Moreover, RT-PCR confirmed the mRNA expression of the hypothetical genes, LOC115950 and MGC4549 (data not shown).

We then performed in situ hybridization to assess the spatial expression of hSM-Calp in adult tissues. This analysis revealed expression of hSM-Calp mRNA in medial SMC of the aorta, visceral SMC of the bladder and stomach and bronchiolar SMC of the lung (Fig. 6). Correct spatial expression of hSM-Calp was also observed in the intestine and uterus whereas the brain, heart, spleen, skeletal muscle, and liver showed only vascular SMC expression (data not shown).

Western blotting studies revealed expression of hSM-Calp protein in such SMC-rich tissues as stomach and uterus (Fig. 7). We exploited the higher affinity of the antisera for human SM-Calp to determine whether
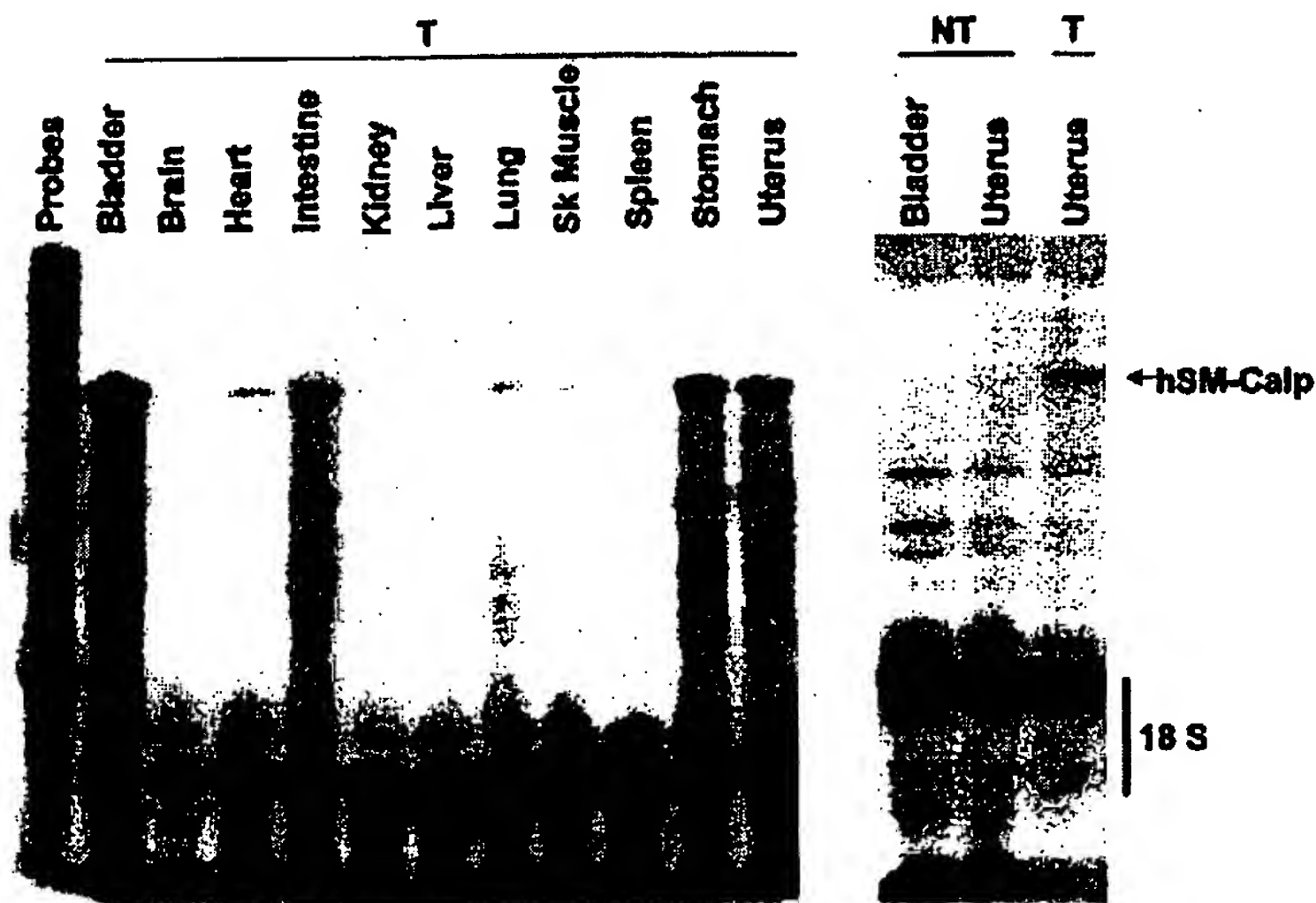

Fig. 4. RNase protection assay of hSM-Calp mRNA expression in adult BAC transgenic mouse tissues. A transgenic mouse line (T) was analyzed for expression of the hSM-Calp transgene and compared with two nontransgenic (NT) tissues (bladder and uterus) using a human-specific riboprobe. The tissue-restricted pattern and level of hSM-Calp expression was identical to that previously reported for the endogenous mouse SM-Calp gene (Fig. 1A in Ref. 31). As expected, no hybridization signal was observed in nontransgenic bladder and uterus (NT). The 18 S signal shows relatively equal total RNA loading across all lanes. Results were reproduced in a second transgenic line.
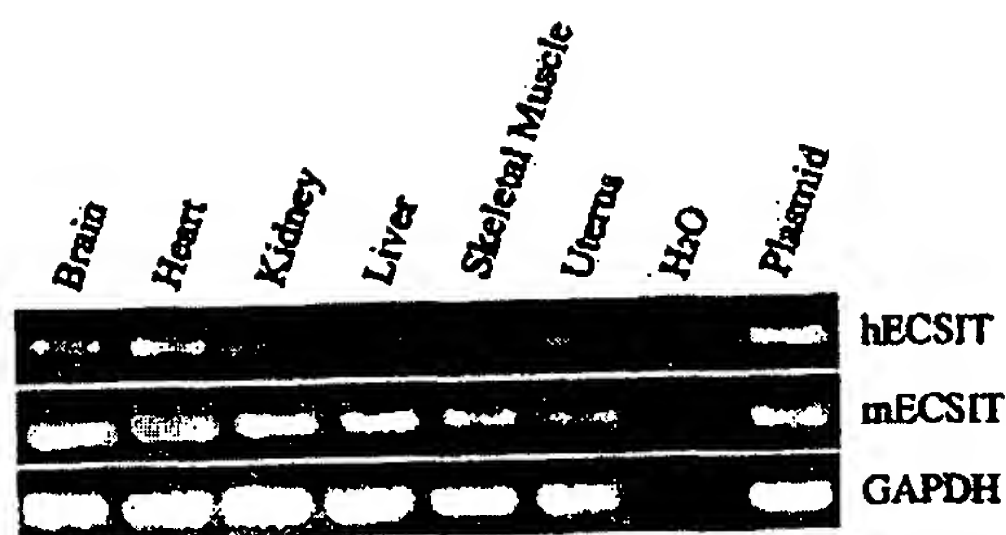
Fig. 5. Human ECSIT (hECSIT) expression in transgenic mouse tissues. Total RNA from various non-SMC tissues was isolated for RT-PCR analysis with the use of primers that distinguish human from mouse ECSIT (mECSIT). Note the widespread expression of both hECSIT and mECSIT.

the hSM-Calp protein was expressed in SMC of these tissues. As shown in Fig. 8, C and D, immunoreactive hSM-Calp was readily detected in both visceral and vascular SMC of the stomach in BAC transgenic mice. Under the conditions employed, little signal was obtained in nontransgenic littermates (Fig. 8, A and B). Collectively, the results from adult BAC transgenic mice indicate that hSM-Calp is expressed in a manner analogous to the endogenous mouse gene at both the mRNA and protein levels.

hSM-Calp expression mirrors endogenous SM-Calp during development. In situ hybridization studies in staged transgenic mouse embryos revealed a pattern of hSM-Calp expression that was very similar to the endogenous SM-Calp gene (31, 48). For example, robust expression of the hSM-Calp transcript was noted in the heart at embryonic (e) 9.5 days with evidence of

Fig. 6. In situ hybridization analysis of hSM-Calp mRNA in adult BAC transgenic mouse tissues. Dark-field microscopy images of transgenic (A,C,E,G) vs. nontransgenic (B,D,F,H) adult tissues processed for in situ hybridization using a human-specific riboprobe. Note the restricted expression of hSM-Calp to vascular SMC of the aorta (A) and visceral SMC of bladder (C), bronchi (E) and stomach (G). Arrows indicate region of visceral SMC in nontransgenic bladder (D) and stomach (G and H). F represents section of a nontransgenic lung. b, Bronchiole.

Stomach Uterus
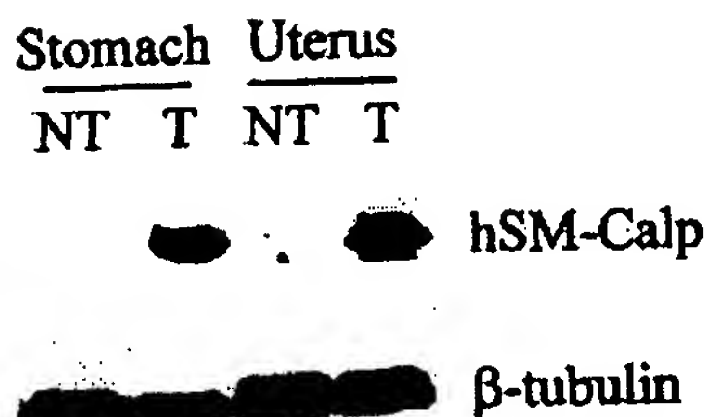
NT T NT T
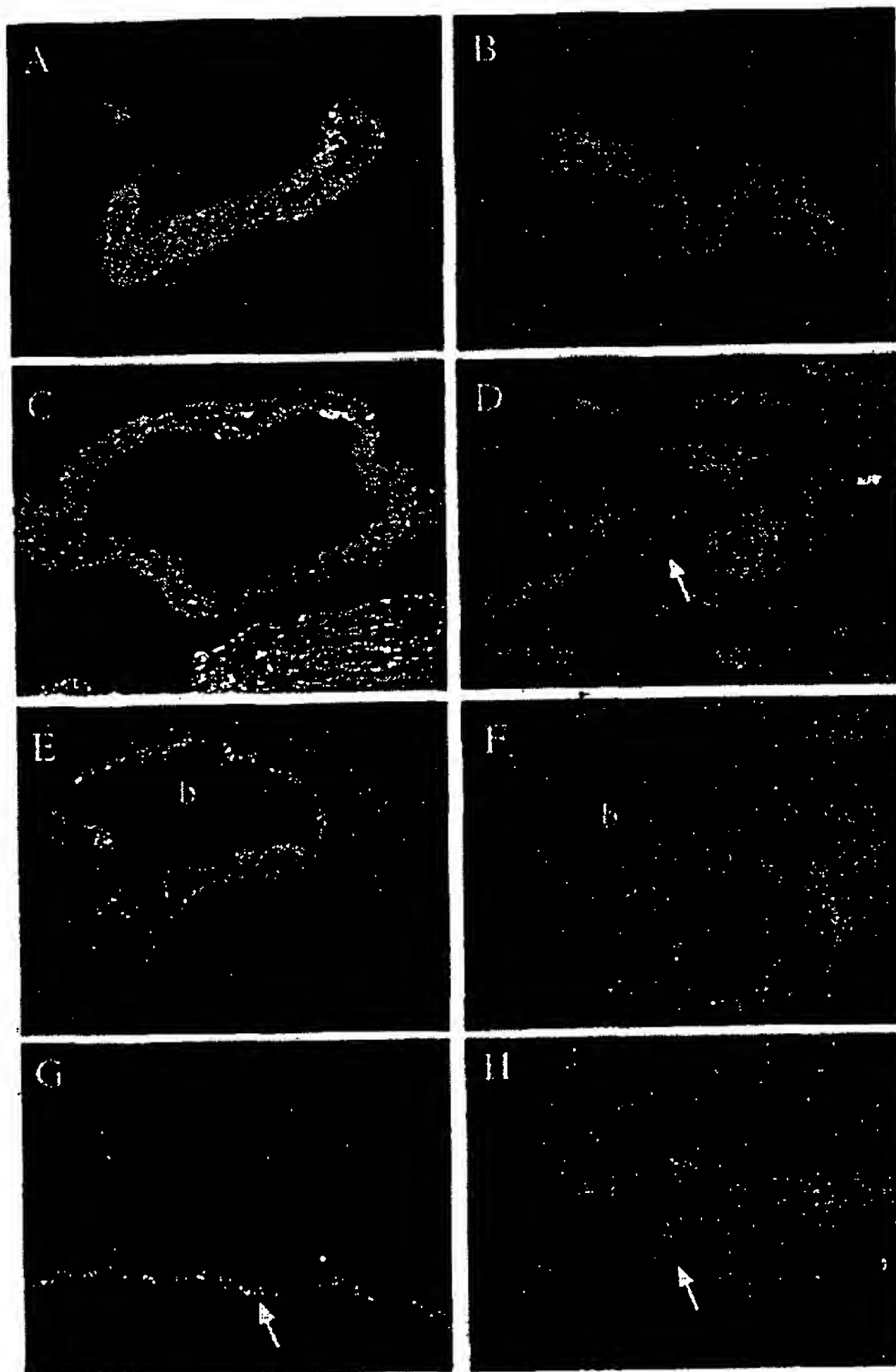
hSM-Calp

β-tubulin

Fig. 7. Western blot analysis of hSM-Calp protein in adult BAC transgenic mouse tissues. Nontransgenic (N) and transgenic (T) stomach and uterus were processed for Western blotting and incubated with antisera either to SM-Calp or β-tubulin. The much higher signal intensity of the transgenic vs. nontransgenic tissues reflects the affinity of the antibody (raised to human SM-Calp) rather than absolute levels. Non-SMC tissues such as liver showed no hSM-Calp expression (data not shown).

embryos clearly exhibit a more intense hybridization signal at all gestational time points. Taken in aggregate, the results of these hSM-Calp transgenic studies show a near complete recapitulation of expression of the endogenous mouse gene.

DISCUSSION

A major innovation of the Human Genome Project has been the development of BACs, which harbor large, stable genomic fragments of DNA (50). BACs have been the workhorses in the sequencing of the human genome (12, 13, 53). In addition, BACs have been used to complement genetic mutations leading to gene identification (2, 43) and to pinpoint distal regulatory elements that confer cell-restricted gene expression in transgenic mice (37). The latter approach has been of particular interest inasmuch as many muscle-restricted transcription units are governed by cis-acting elements that may reside tens of kilobases away from the core promoter region (9, 47, 57). Defining such distal elements through small phage genomic clones can be time consuming and may have untoward effects on the expression of the reporter transgene (see below). In this report we used a BAC harboring the hSM-Calp transcription unit to show first that the expression of hSM-Calp exhibited similar regulatory expression in an in vitro model of cellular differentiation wherein expression of SM-Calp is attenuated on serum withdrawal and differentiation (33). The same BAC was then integrated into the mouse genome to assess hSM-Calp expression in an in vivo context where the trans-

dorsal aortic SMC expression (Fig. 9, A–D). Intense heart and dorsal aortic expression continued through e12.5 days (Fig. 9, E–H). Widespread visceral SMC expression of the hSM-Calp transgene was observed in the bladder, gut, esophagus, and bronchi at e14.5-e15.5 days (Fig. 9, I–P). Interestingly, expression of hSM-Calp mRNA in the heart occurs as late as e15.5 days, a time in which the expression of the endogenous mouse SM-Calp gene has all but vanished in this tissue (Fig. 9, M–P) (31). Although we used a riboprobe to the 3′ untranslated region of hSM-Calp (77% homology to mouse), some cross-hybridization to the endogenous mouse SM-Calp mRNA was observed in nontransgenic littermates (Fig. 9, G and K). However, the transgenic
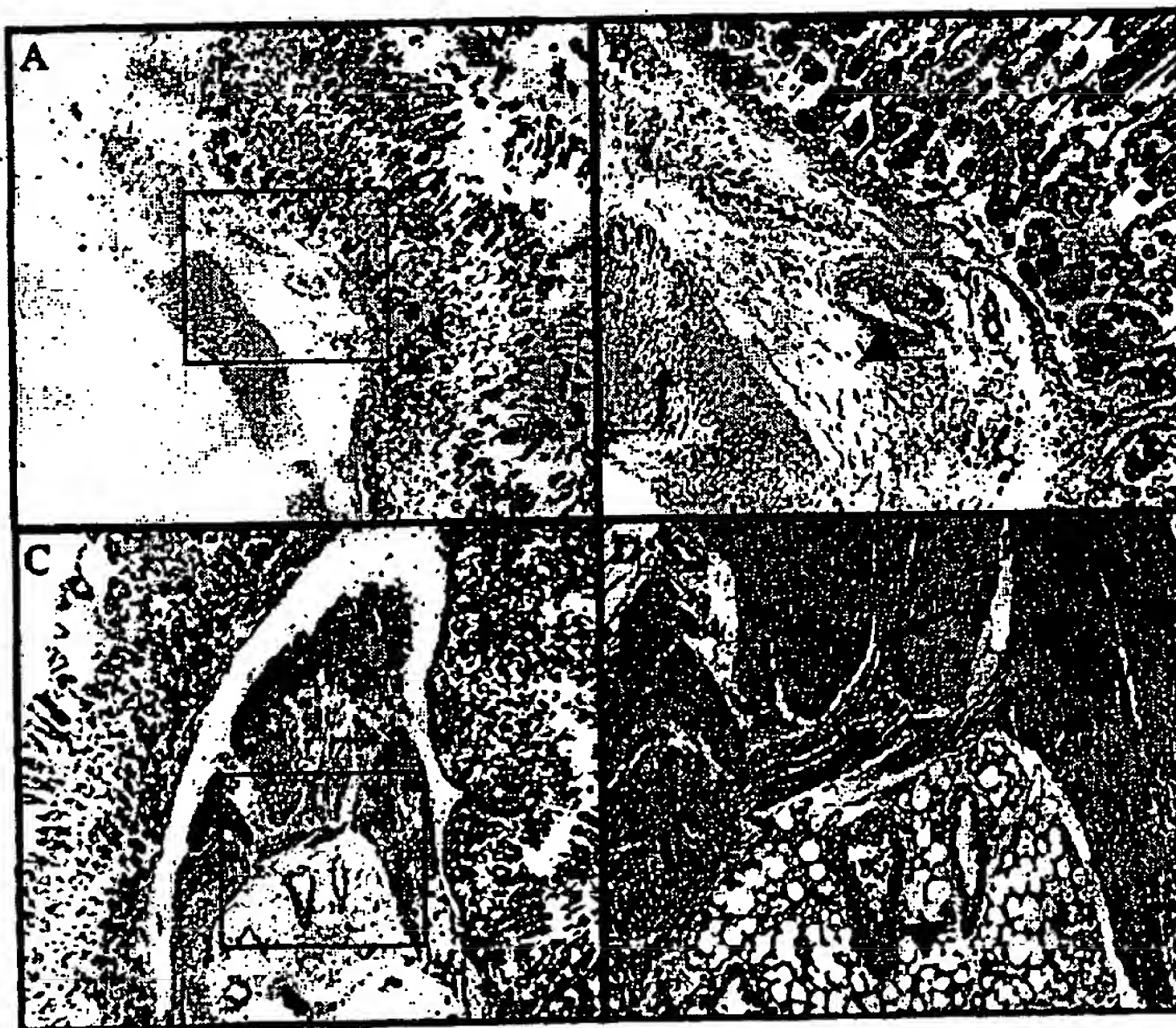

Fig. 8. Immunolocalization of hSM-Calp in vascular and visceral SMC of the stomach. Nontransgenic (A and B) and transgenic (C and D) stomach tissues were processed for immunohistochemistry with 3-amino-9-ethylcarbazole as substrate (red color). Note that under the conditions used, the antibody to SM-Calp preferentially detects the hSM-Calp protein in both vascular (▲) and visceral (arrow) SMC of the BAC transgenic stomach (compare B with D). Magnifications were ×40 (A and C) and ×100 (B and D).
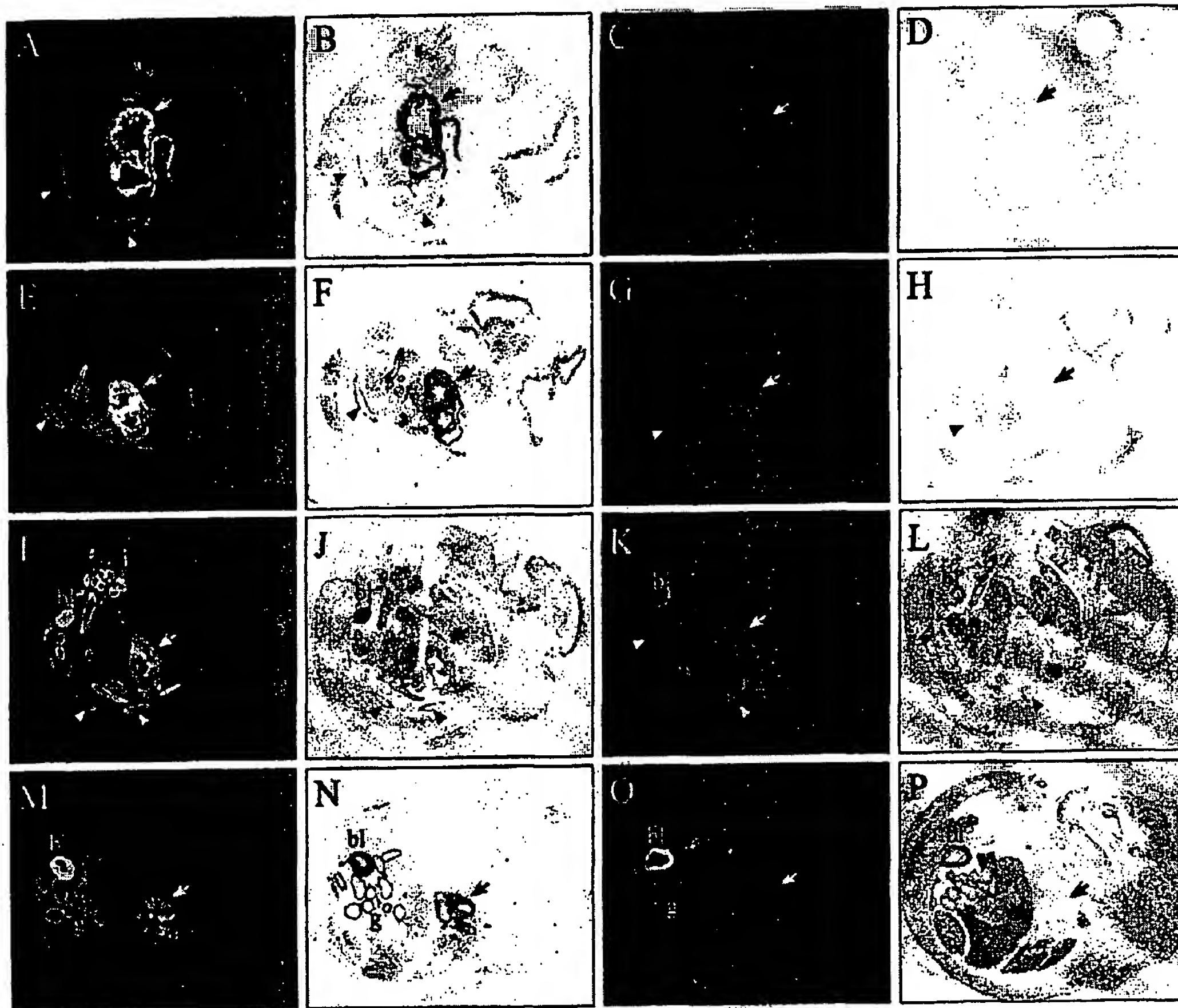
Fig. 9. In situ hybridization of hSM-Calp mRNA in developmentally staged mouse embryos. Saggital sections from e9.5 (A–D), e12.5 (E–H), e14.5 (I–L), or e15.5 (M–P) day embryos of transgenic (A, B, E, F, I, J, M, and N) or nontransgenic (C, D, G, H, K, L, O, and P) origin were hybridized to an antisense hSM-Calp riboprobe and processed for darkfield (A, C, E, G, I, K, M, and O) or brightfield (B, D, F, H, J, L, N, and P) microscopy. hSM-Calp is robustly expressed in embryonic heart (arrows) as early as e9.5 days (A and B) and persists through e15.5 days (M and N), a time at which the endogenous mouse SM-Calp gene is turned off in the heart (O and P) and (31). High-level hSM-Calp expression is also observed in the dorsal aorta (arrowheads in A, B, E, F, I, and J), bladder and gut (I, J, M, and N), and esophagus (I and J). For unclear reasons, cross-hybridization of the hSM-Calp riboprobe to endogenous mouse SM-Calp was evident in some of the nontransgenic embryos (heart in G and H as well as bladder and gut in K, L, O, and P). bl, Bladder; es, esophagus; g, gut. See text for more details.

gene would be subject to the complex signaling cues that normally converge on a transcription unit within highly ordered chromatin. Expression of the hSM-Calp transgene was restricted to both vascular and visceral SMC of adult transgenic mice, a result that contrasts with SM22 and CRP1 promoter mice, which only expressed a lacZ reporter gene in arterial SMC of adult mice (15, 21, 23, 34). During development, the expression of hSM-Calp mirrored that of the endogenous mouse gene with one notable exception; cardiac expression of hSM-Calp appeared to persist longer than the endogenous mouse gene (Fig. 9) and (31). Consistent with endogenous data (31, 48), hSM-Calp is not expressed in developing somites, a finding that contrasts with all other SMC markers known to be expressed in sarcomeric muscle. It will be of interest to determine the transcriptional mechanisms for such somitic silencing as well as the protracted expression of hSM-Calp in developing cardiac muscle.

How might one begin to define critical regulatory elements governing the complete expression profile of hSM-Calp from a >100-kb BAC? One approach would be to serially delete the BAC clone, a process that could identify distal elements within intergenic regions or even neighboring transcription units as appears to be the case with the skeletal muscle transcription factor

MRF4 (57). Here we made an aggressive deletion of the BAC and showed that proper hSM-Calp expression remained in the BC3H1 cell line. The BC3H1 model system of SM-Calp gene regulation therefore represents a fast and easy system to identify candidate deletion clones that would ultimately require testing in vivo. Another approach of locating discrete cis elements in artificial chromosomes involves the random cloning and testing of BAC fragments in an enhancer trap assay. Studies to date using this approach, however, have failed to identify SMC-specific enhancers in the BAC containing hSM-Calp. An alternative approach that is increasingly emerging as a powerful tool to uncover important cis elements in the genome is through "phylogenetic fingerprinting," in which comparative sequence analysis is undertaken between two species (e.g., mouse vs. human) (54). We previously used such an analysis to show the existence of several evolutionarily conserved CArG boxes in the first intron of SM-Calp, which appear to confer SMC-specific enhancer activity in vitro (30). An important issue to investigate therefore will be whether these intronic CArG boxes direct SMC-specific expression of the hSM-Calp transgene in vivo. Such point mutation experiments are possible using methods of homologous recombination (35).

Although the use of BACs for purposes of defining regulatory elements may appear to be a somewhat daunting endeavor, we favor this approach over conventional lacZ studies for several reasons. First, regulatory elements may reside considerable distances from the core promoter, making conventional genomic phage approaches both time and labor intensive. Experience with such muscle-restricted promoters as SM22 (15, 21, 34), CRP1 (23), Csx/Nkx2.5 (52), MLC1F/3F (7), myf-5 (41), and desmin (18), to name a few, indicate the multiplicity of regulatory cassettes in and around gene loci. Artificial chromosomes offer the opportunity to potentially capture all regulatory elements of a gene, thus simplifying the analysis of gene regulation. This is of particular importance, given the fact that distal elements are likely to interact with one another and the core promoter to facilitate proper gene transcription. Second, incorporation of an artificial reporter gene such as lacZ downstream of promoter sequences alters the native sequence landscape of a gene and may thus disrupt important spatial or sequence configurations that must be preserved for proper gene transcription (10). Alternatively, epigenetic alterations (e.g., methylation) in the promoter sequences or reporter itself could compromise reporter readout and hence confound interpretations of true promoter activity (6, 38). In this context, we have recently observed transgenic mice carrying a 1.4-kb SM22 promoter fused to lacZ to express the reporter appropriately during development only to lose expression entirely during postnatal development (unpublished data, J. M. Miano). These artifacts of reporter gene expression are greatly minimized when the reporter is the natural transcription unit itself in the context of a large genomic landscape. Ultimately, definitive proof of the importance of a regulatory element in governing proper gene expression will require gene-targeted ablation studies (39).

Another major advantage of BACs is that they often contain more than one transcription unit, which provides important intergenic sequence information. There is a growing consensus that intergenic sequences harbor important regulatory elements that facilitate proper gene transcription. How, for example, can SM-Calp be restricted to SMC lineages when it is less than 10 kb away from the widely expressed ECSIT transcription unit? One way of attaining this important control is through intergenic boundary elements (e.g., insulators) that serve to shield a locus from the effects of an adjacent gene's promiscuous transcriptional apparatus (3). In silico analyses have failed to reveal the presence of a consensus enhancer-blocking boundary element (3) in the intergenic region between SM-Calp and ECSIT. It is possible, however, that other boundary elements exist in this region that will require more extensive computer and wet-lab analyses to uncover.

In addition to showing SMC-restricted hSM-Calp mRNA expression, we provide evidence for both vascular and visceral SMC expression of the human SM-Calp protein. There are only 8 amino acid substitutions over the 297 amino acid primary sequences of mouse and human SM-Calp (6 are conservative). Thus it is reasonable to assume that the human SM-Calp protein will function in an identical manner as the endogenous mouse protein. These BAC transgenic mice therefore will be useful to breed to homozygosity and assess the physiological effects of SM-Calp over-expression in vivo. Moreover, the present BAC could be used to rescue the bone and vascular phenotypes observed in the SM-Calp null mice (29, 56). Interpreting such studies, however, could be confounded by the presence of other human transgenes on the BAC (e.g., hECSIT).

In summary, we have used a BAC harboring the hSM-Calp locus to show, for the first time, the complete recapitulation of the endogenous SM-Calp's promoter activity in developing and postnatal transgenic mouse tissues. Future studies will examine the importance of known (intronic CArG elements) and as yet to be defined regulatory elements in conferring SMC-restricted expression of hSM-Calp in the context of the present BAC. These and other studies will be instrumental in gaining a complete fingerprint of the transcriptional program underlying SMC differentiation in vivo and how this program of differentiation is perturbed in disease states.

REFERENCES

1. Ainscough JFX, John RM, Barton SC, and Surani MA. A skeletal muscle-specific mouse *Igf2* repressor lies 40 kb downstream of the gene. *Development* 127: 3923–3930, 2000.

2. Antoch MP, Song EJ, Chang AM, Vitaterna MH, Zhao Y, Wilsbacher LD, Sangoram AM, King DP, Pinto LH, and Takahashi JS. Functional identification of the mouse circadian clock gene by transgenic BAC rescue. *Cell* 89: 655–667, 1997.

3. Bell AC, West AG, and Felsenfeld G. Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science* 291: 447–450, 2001.

4. Burke DT and Olson MV. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 236: 806–812, 1987.

5. Chomczynski P and Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162: 156–159, 1987.

6. Cohen-Tannoudji M, Vandormael S, Drezen JM, Mercier P, Babinet C, and Morello D. LacZ sequences prevent regulated expression of housekeeping genes. *Mech Dev* 90: 29–39, 2000.

7. Franco D, Kelly R, Lamers WH, Buckingham M, and Moorman AFM. Regionalized transcriptional domains of myosin light chain 3f transgenes in the embryonic mouse heart: morphogenetic implications. *Dev Biol* 188: 17–33, 1997.

8. Franz WM, Mueller OJ, Fleischmann M, Babij P, Frey N, Mueller M, Besenfelder U, Moorman AFM, Brem G, and Katus HA. The 2.3 kb smooth muscle myosin heavy chain promoter directs gene expression into the vascular system of transgenic mice and rabbits. *Cardiovasc Res* 43: 1040–1048, 1999.

9. Goldhamer DJ, Faerman A, Shani M, and Emerson CP Jr. Regulatory elements that control the lineage-specific expression of myoD. *Science* 256: 538–542, 1992.

10. Guy LG, Kothary R, DeRepentigny Y, Delvoye N, Ellis J, and Wall L. The β-globin locus control region enhances transcription of but does not confer position-independent expression onto the lacZ gene in transgenic mice. *EMBO J* 15: 3713–3721, 1996.

11. Herring BP and Smith AF. Telokin expression is mediated by a smooth muscle cell-specific promoter. *Am J Physiol Cell Physiol* 270: C1656–C1665, 1996.

12. International Human Genome Mapping Consortium. A physical map of the human genome. *Nature* 409: 934–941, 2001.

13. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 409: 860–921, 2001.

14. Kim S, Lin H, Barr E, Chu L, Leiden JM, and Parmacek MS. Transcriptional targeting of replication-defective adenovirus transgene expression to smooth muscle cells in vivo. *J Clin Invest* 100: 1006–1014, 1997.

15. Kim S, Ip HS, Lu MM, Clendenin C, and Parmacek MS. A serum response factor-dependent transcriptional regulatory program identifies distinct smooth muscle cell sublineages. *Mol Cell Biol* 17: 2266–2278, 1997.

16. Kitami Y, Maguchi M, Nishida W, Okura T, Kohara K, and Hiwada K. The unique 5-flanking region of the human basic calponin gene. *Hypertens Res* 22: 187–193, 1999.

17. Kopp E, Medzhitov R, Carothers J, Xiao C, Douglas I, Janeway CA, and Ghosh S. ECSIT is an evolutionarily conserved intermediate in the Toll/IL-1 signal transduction pathway. *Genes Dev* 13: 2059–2071, 1999.

18. Kuisk IR, Li H, Tran D, and Capetanaki Y. A single MEF2 site governs desmin transcription in both heart and skeletal muscle during mouse embryogenesis. *Dev Biol* 174: 1–13, 1996.

19. Lakshmanan G, Lieuw KH, Grosveld F, and Engel JD. Partial rescue of GATA-3 by yeast artificial chromosome transgenes. *Dev Biol* 204: 451–463, 1998.

20. Li L, Miano JM, Cserjesi P, and Olson EN. SM22α, a marker of adult smooth muscle, is expressed in multiple myogenic lineages during embryogenesis. *Circ Res* 78: 188–195, 1996.

21. Li L, Miano JM, Mercer B, and Olson EN. Expression of the SM22α promoter in transgenic mice provides evidence for distinct transcriptional regulatory programs in vascular and visceral smooth muscle cells. *J Cell Biol* 132: 849–859, 1996.

22. Li L, Liu ZC, Mercer B, Overbeek P, and Olson EN. Evidence for serum response factor-mediated regulatory networks governing SM22α transcription in smooth, skeletal, and cardiac muscle cells. *Dev Biol* 187: 311–321, 1997.

23. Lilly B, Olson EN, and Beckerle MC. Identification of a CArG box-dependent enhancer within the cysteine-rich protein 1 gene that directs expression in arterial but not venous or visceral smooth muscle cells. *Dev Biol* 240: 531–547, 2001.

24. Ma WJ, Cheng S, Campbell C, Wright A, and Furneaux H. Cloning and characterization of HuR, a ubiquitously expressed Elav-like protein. *J Biol Chem* 271: 8144–8151, 1996.

25. Mack CP and Owens GK. Regulation of smooth muscle α-actin expression in vivo is dependent on CArG elements within the 5′ and first intron promoter regions. *Circ Res* 84: 852–861, 1999.

26. Madsen CS, Regan CP, Hungerford JE, White SL, Manabe I, and Owens GK. Smooth muscle-specific expression of the smooth muscle myosin heavy chain gene in transgenic mice requires 5′-flanking and first intronic DNA sequence. *Circ Res* 82: 908–917, 1998.

27. Maguchi M, Nishida W, Kohara K, Kuwano A, Kondo I, and Hiwada K. Molecular cloning and gene mapping of human basic and acidic calponins. *Biochem Biophys Res Commun* 217: 238–244, 1995.

28. Manabe I and Owens GK. CArG elements control smooth muscle subtype-specific expression of smooth muscle myosin in vivo. *J Clin Invest* 107: 823–834, 2001.

29. Matthew JD, Khromov AS, McDuffie MJ, Somlyo AV, Somlyo AP, Taniguchi S, and Takahashi K. Contractile properties and proteins of smooth muscles of a calponin knockout mouse. *J Physiol (Lond)* 529: 811–824, 2000.

30. Miano JM, Carlson MJ, Spencer JA, and Misra RP. Serum response factor-dependent regulation of the smooth muscle calponin gene. *J Biol Chem* 275: 9814–9822, 2000.

31. Miano JM and Olson EN. Expression of the smooth muscle cell calponin gene marks the early cardiac and smooth muscle cell lineages during mouse embryogenesis. *J Biol Chem* 271: 7095–7103, 1996.

32. Miano JM, Krahe R, Garcia E, Elliott JM, and Olson EN. Expression, genomic structure and high resolution mapping to 19p13.2 of the human smooth muscle cell calponin gene. *Gene* 197: 215–224, 1997.

33. Miano JM, Thomas S, and Disteche CM. Expression and chromosomal mapping of the mouse smooth muscle calponin gene. *Mamm Genome* 12: 187–191, 2001.

34. Moessler H, Mericskay M, Li Z, Nagl S, Paulin D, and Small JV. The SM 22 promoter directs tissue-specific expression in arterial but not in venous or visceral smooth muscle cells in transgenic mice. *Development* 122: 2415–2425, 1996.

35. Muyrers JPP, Zhang Y, Benes V, Testa G, Ansorge W, and Stewart AF. Point mutation of bacterial artificial chromosomes by ET recombination. *EMBO J* 1: 239–243, 2000.

36. Nielsen LB, McCormick SPA, Pierotti V, Tam C, Gunn MD, Shizuya H, and Young SG. Human apolipoprotein B transgenic mice generated with 207- and 145-kilobase pair bacterial artificial chromosomes: evidence that a distant 5′ element confers appropriate transgene expression in the intestine. *J Biol Chem* 272: 29752–29758, 1997.

37. Nielsen LB, Kahn D, Duell T, Weier HUG, Taylor S, and Young SG. Apolipoprotein B gene expression in a series of human apolipoprotein B transgenic mice generated with recA-assisted restriction endonuclease cleavage-modified bacterial artificial chromosomes: an intestine-specific enhancer element is located between 54 and 62 kilobases 5′ to the structural gene. *J Biol Chem* 273: 21800–21807, 1998.

38. Nilsson E and Lendahl U. Transient expression of a human β-actin promoter/lacZ gene introduced into mouse embryos correlates with a low degree of methylation. *Mol Reprod Dev* 34: 149–157, 1993.

39. Oosthuyse B, Moons L, Storkebaum E, Beck H, Nuyens D, Brusselmans K, Van Dorpe J, Hellings P, Gorselink M, Heymans S, Theilmeier G, Dewerchin M, Laudenbach V, Vermylen P, Raat H, Acker T, Vleminckx V, Van Den

Bosch L, Cashman N, Fujisawa H, Drost MR, Sciot R, Bruyninckx F, Hicklin DJ, Ince C, Gressens P, Lupu F, Plate KH, Robberecht W, Herbert JM, Collen D, and Carmeliet P. Deletion of the hypoxia-response element in the vascular endothelial growth factor promoter causes motor neuron degeneration. Nat Genet 28: 131–138, 2001.

40. Owens GK. Regulation of differentiation of vascular smooth muscle cells. Physiol Rev 75: 487–517, 1995.

41. Patapoutian A, Miner JH, Lyons GE, and Wold B. Isolated sequences from the linked Myf-5 and MRF4 genes drive distinct patterns of muscle-specific expression in transgenic mice. Development 118: 61–69, 1993.

42. Peterson KR, Clegg CH, Huxley C, Josephson BM, Haugen HS, Furukawa T, and Stamatoyannopoulos G. Transgenic mice containing a 248-kb yeast artificial chromosome carrying the human β-globin locus display proper developmental control of human globin genes. Proc Natl Acad Sci USA 90: 7593–7597, 1993.

43. Probst FJ, Fridell RA, Raphael Y, Saunders TL, Wang A, Liang Y, Morell RJ, Touchman JW, Lyons RH, Noben-Trauth K, Friedman TB, and Camper SA. Correction of deafness in shaker-2 mice by an unconventional myosin in a BAC transgene. Science 280: 1444–1447, 1998.

44. Qian J, Kumar A, Szucsik JC, and Lessard JL. Tissue and developmental specific expression of murine smooth muscle gamma-actin fusion genes in transgenic mice. Dev Dyn 207: 135–144, 1996.

45. Raguz S, Hobbs C, Yague E, Ioannou PA, Walsh FS, and Antoniou M. Muscle-specific locus control region activity associated with the human desmin gene. Dev Biol 201: 26–42, 1998.

46. Regan CP, Adam PJ, Madsen CS, and Owens GK. Molecular mechanisms of decreased smooth muscle differentiation marker expression after vascular injury. J Clin Invest 106: 1139–1147, 2000.

47. Rosenthal N, Kornhauser JM, Donoghue M, Rosen KM, and Merlie JP. Myosin light chain enhancer activates muscle-specific, developmentally regulated gene expression in transgenic mice. Proc Natl Acad Sci USA 86: 7780–7784, 1989.

48. Samaha FF, Ip HS, Morrisey EE, Seltzer J, Tang Z, Solway J, and Parmacek MS. Developmental pattern of expression and genomic organization of the calponin-h1 gene: a contractile smooth muscle cell marker. J Biol Chem 271: 395–403, 1996.

49. Schedl A, Montoliu L, Kelsey G, and Schütz G. A yeast artificial chromosome covering the tyrosinase gene confers copy number-dependent expression in transgenic mice. Nature 362: 258–261, 1993.

50. Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, Tachiiri Y, and Simon M. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. Proc Natl Acad Sci USA 89: 8794–8797, 1992.

51. Sinn PL, Davis DR, and Sigmund CD. Highly regulated cell type-restricted expression of human renin in mice containing 140- or 160-kilobase pair P1 phage artificial chromosome transgenes. J Biol Chem 274: 35785–35793, 2001.

52. Tanaka M, Burns-Wechsler S, Lee IW, Yamasaki N, Lawitts JA, and Izumo S. Complex nuclear cis-acting elements regulate expression of the cardiac specifying homeobox gene Csx/Nkx2.5. Development 126: 1439–1450, 1999.

53. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, et al. The sequence of the human genome. Science 291: 1304–1351, 2001.

54. Wasserman WW, Palumbo M, Thompson W, Fickett JW, and Lawrence CE. Human-mouse genome comparisons to locate regulatory sites. Nat Genet 26: 225–228, 2000.

55. Yet SF, Folta SC, Jain MK, Hsieh CM, Maemura K, Layne MD, Zhang D, Marria PB, Yoshizumi M, Chin MT, Perrella MA, and Lee ME. Molecular cloning, characterization, and promoter analysis of the mouse Crp2/SMLim gene: preferential expression of its promoter in the vascular smooth muscle cells of transgenic mice. J Biol Chem 273: 10530–10537, 1998.

56. Yoshikawa H, Taniguchi S, Yamamura H, Mori S, Sugimoto M, Miyado K, Nakamura K, Nakao K, Katsuki M, Shibata N, and Takahashi K. Mice lacking smooth muscle calponin display increased bone formation that is associated with enhancement of bone morphogenetic protein responses. Genes to Cells 3: 685–695, 1998.

57. Zweigerdt R, Braun T, and Arnold HH. Faithful expression of the myf-5 gene during mouse myogenesis requires distant control regions: a transgene approach using yeast artificial chromosomes. Dev Biol 192: 172–180, 1997.